

Einstellung zur Hamburger Innenstadt: eine Auswertung mit Hilfe der Korrespondenzanalyse

Blasius, Jörg

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Blasius, J. (1987). Einstellung zur Hamburger Innenstadt: eine Auswertung mit Hilfe der Korrespondenzanalyse. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 21, 29-51. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-205249>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.



**Einstellung zur Hamburger Innenstadt
Eine Auswertung mit Hilfe der Korrespondenzanalyse
von Jörg Blasius**

Abstract

The purpose of the article is to introduce the method of correspondence analysis, using data from a survey on attitudes toward the Hamburg downtown. In 1984, a telephone survey was conducted to assess the attitudes of 832 Hamburg residents toward the downtown and urban sub-centers. In addition, the spatial distribution of their leisure activities was included in the questionnaire. The results of the study indicate that there is nothing like a "typical downtown visitor", however, there are subpopulations, varying by their orientation towards the local center and the downtown. Correspondence analysis provides an adequate method to analyze and describe these subpopulations. Four groups of center users emerge from this analysis.

Zusammenfassung

Dieser Bericht soll eine Einführung in das Verfahren der Korrespondenzanalyse geben, wofür Daten über die Einstellung zur Hamburger Innenstadt und den Stadtteilen verwendet wurden. 1984 fand im Rahmen eines Forschungsprojektes eine telefonische Befragung von 832 Hamburgern statt. Ziel der Studie war es, in Erfahrung zu bringen, ob und von wem die Hamburger Innenstadt als innerstädtischer Mittelpunkt angesehen wird, ferner, ob einzelne Freizeitaktivitäten eher in der Innenstadt oder einem Stadtteilzentrum ausgeübt werden. Um eine Typologie der Befragten zu entwickeln, wurde als multivariates Auswertungsverfahren die Korrespondenzanalyse verwendet. Der Aufsatz stellt sowohl das Verfahren als auch inhaltliche Ergebnisse dar. Vier Gruppen von Stadtbenutzern lassen sich unterscheiden. Es zeigte sich, daß es den typischen Innenstadtbesucher in Hamburg nicht gibt, demgegenüber aber Subpopulationen bestehen, die verstärkt den eigenen Stadtteil aufsuchen. Der verwendete Datensatz wird z.Zt. im Zentralarchiv archiviert.



1. Einleitung

In den Sozialwissenschaften werden Verfahren benötigt, um komplexe Zusammenhänge zwischen Daten multivariat auswerten zu können. Daß hierbei nicht immer statistisch sauber gearbeitet wird und als Rechtfertigung für dieses Vorgehen andere Autoren zitiert werden, die in ähnlicher Weise vorgegangen sind, bedarf wohl keiner Belege. So werden Faktoren-, Pfad- oder Diskriminanzanalysen gerechnet, obgleich keine metrischen, sondern nur ordinale oder gar nur nominale Daten vorliegen. Gerade in der Soziologie verfügen wir relativ selten über kontinuierliche Variablen, die klassischen unabhängigen Variablen sind vielmehr sozio-demografische Merkmale der Person.

2. Zur Korrespondenzanalyse

Sollen Auswertungen einerseits nicht auf der Anwendung adäquater Verfahren basieren, andererseits aber auch nicht bei Kreuztabellen aufhören, so werden multivariate Verfahren benötigt, die keine Voraussetzungen bezüglich des Skalenniveaus machen. Eines dieser Verfahren ist die in Frankreich (BENZECRI, 1963; ESCOFIER-CORDIER, 1965) entwickelte Korrespondenzanalyse. Ihre ersten Ansätze können auf HIRSCHFELD (1935) zurückgeführt werden, der eine algebraische Formulierung für die "Korrelation" von Zeilen und Spalten einer Kontingenztafel gab. GUTTMAN (1941, 1950) entwickelte dies als Skalierungsverfahren mit Hilfe der Hauptkomponentenanalyse weiter, ebenso wie HAYASHI (1950, 1952, 1954) mit seiner "quantification of qualitative data".

Ähnlich der Korrespondenzanalyse sind die Verfahren des "reciprocal averaging" (insb. HILL, 1973) und des "dual or optimal scaling" (insb. NISHISATO, 1978, 1980). Die zur Zeit ausführlichste Beschreibung und die statistischen Grundlagen der Korrespondenzanalyse liefert das Buch von GREENACRE (1984). Eine kurze Einführung in das Verfahren anhand eines historischen Datensatzes gibt die Arbeit von BLASIUS (1987).

Bei der Korrespondenzanalyse handelt es sich primär um ein Verfahren zur grafischen Darstellung der Spalten und Zeilen von zweidimensionalen (two-way) Kontingenztabellen (GREENACRE, 1981, S. 119). Wie bei der Hauptkomponentenanalyse gibt es einen Satz von orthogonal aufeinander stehenden Vektoren, die einen niederdimensionalen Raum aufspannen, im Fall einer



beabsichtigten grafischen Darstellung höchstens zwei. (Es lassen sich prinzipiell auch höherdimensionale Räume einführen. Diese können aber in Analogie zu grafischen Rotationsverfahren der Faktorenanalyse nur ebenenweise ausgeplottet werden.) Neu im Sinne der praktischen Anwendung ist die verwendete Metrik. Hier unterscheidet sich die Korrespondenzanalyse von Verfahren wie der multidimensionalen Skalierung oder der Clusteranalyse.

Mit der Korrespondenzanalyse lassen sich sowohl Daten auf Individualebene als auch auf Aggregatenebene auswerten, wobei in dieser Studie nur letzteres vorgestellt werden soll. Eine Anwendung mit Individualdaten geben DANGSCHAT und BLASIUS (1987) in ihrer Beschreibung von Warschau. Zum leichteren Verständnis sei ein Rechenbeispiel für den bivariaten Fall vorgeführt, welches später auf die multivariate Anwendung erweitert wird.

3. Die Daten

Im Rahmen einer vergleichenden Studie über die Innenstädte von Hamburg und Baltimore wurde eine telefonische Befragung einer Wahrscheinlichkeitsstichprobe von 832 Hamburgern im Frühjahr 1984 durchgeführt. Ziel dieses Projektes war es, die relative ökonomische Position der Innenstadt gegenüber den Stadtteilzentren zu ermitteln (vgl. FRIEDRICHS und GOODMAN, 1987). Eine Fragestellung war in diesem Zusammenhang, welche der acht Freizeitaktivitäten von welchen Personengruppen wie oft in der Hamburger Innenstadt ("City") bzw. dem eigenen Stadtteilzentrum oder an anderer Stelle in der Stadt ausgeübt wird. Die folgenden Ausführungen beschränken sich auf die Alternative Innenstadt vs. Stadtteilzentrum.

Hierbei interessierte insbesondere, ob es eine Gruppe von Personen (die wir im folgenden kurz "Typ" nennen) gibt, die relativ viele Aktivitäten in der Innenstadt ausübt, eine weitere, die relativ viele Aktivitäten im eigenen Stadtteil ausübt, ferner eine, die relativ viele Aktivitäten sowohl in der Innenstadt als auch im eigenen Stadtteil ausübt, und schließlich eine, die relativ wenige Aktivitäten in den genannten Gebieten ausübt. Forschungsziel war es, eine Art Diskriminanzfunktion zu finden, die diese vier Gruppen eindeutig trennt und ihnen charakteristische Eigenschaften zuordnet.

Die Aktivitäten wurden operationalisiert durch je acht Einzelaktivitäten (für die Innenstadt und den eigenen Stadtteil getrennt; Besuch von: Restaurants,



Kneipen, Konzerten, Kino, Sportveranstaltungen, kirchlichen Veranstaltungen, Museen und Oper), die zu einer Aktivitätsskala zusammengefaßt wurden. Je mehr dieser Aktivitäten der Befragte im vorausgegangenen Jahr mindestens einmal ausgeübt hat, desto höher sein Wert auf der jeweiligen Aktivitätsskala. Beide Aktivitätsskalen wurden über den Median gesplittet. Dabei wurde, um die Differenz der Werte (1,66 für den eigenen Stadtteil vs. 2,47 für die City) zu berücksichtigen, in beiden Fällen abgerundet. Die Verteilung der Aktivitäten und der sich daraus ableitenden vier Typen zeigt Tabelle 1.

Tabelle 1: Verteilung der Anzahl der Aktivitäten auf die Innenstadt und den eigenen Stadtteil

*)	Anzahl der Aktivitäten in der Innenstadt							Summe
	0	1	2	3	4	5	6-8	
0	38	28	23	22	14	11	5	141
1	52	50	45	38	34	15	7	241
2	22	40	32	45	43	20	8	210
3	6	11	28	19	21	21	9	115
4	1	9	16	17	16	12	4	75
5	3	6	6	6	3	7	4	35
6-8	0	2	2	1	5	2	3	15
Summe	122	146	152	148	136	88	40	832

Chi Square = 105,0; $p < .001$; Tau b = .22

*) Anzahl der Aktivitäten im eigenen Stadtteil

Durch das Splitten über die Mediane ergibt sich u.a. eine Gruppe von 146 Personen (Summe im rechten oberen Quadranten von Tabelle 1), die im folgenden als "Citytyp" bezeichnet werden sollen. Ihr Merkmal ist es, nicht mehr als eine der genannten Aktivitäten im eigenen Stadtteil auszuüben, hingegen mindestens drei in der Innenstadt.

Wie aus Tabelle 1 ferner ersichtlich wird, ist die mittlere Anzahl der Aktivitäten, die in der Innenstadt ausgeübt werden, wesentlich höher als im eigenen Stadtteil. Dies läßt sich darauf zurückführen, daß hier implizit auch die Ausstattung des Stadtteils mitgemessen wird, da die Subzentren in der Regel nicht über Kino, Theater, Museen oder gar über eine Oper



verfügen. Personen, die auf ein derartiges Vergnügen nicht verzichten wollen, sind somit darauf angewiesen, ihren eigenen Stadtteil zu verlassen und ein besser ausgestattetes Zentrum bzw. die Innenstadt aufzusuchen. Wenn hier gerade die Innenstadt die größte Anziehung ausübt, so deshalb, weil sie sämtliche Freizeitmöglichkeiten. (insb. kulturelle) bietet.

Eine der zu vermutenden unabhängigen Variablen ist das Alter der Befragten. Es wurde in sechs Gruppen zusammengefaßt, sodann mit den vier Typen kreuztabelliert. Die Verteilung nach Alter und Typenzugehörigkeit ist in Tabelle 2 dargestellt.

Tabelle 2: Verteilung der Typen auf die Altersklassen

	macht nichts	nur eig. Stadtteil	nur City	Stadtteil und City	Summe
bis 25 Jahre	15	23	17	41	96
25 bis 34 Jahre	35	37	32	66	170
35 bis 44 Jahre	29	46	33	58	166
45 bis 54 Jahre	37	34	20	48	139
55 bis 64 Jahre	47	20	23	25	115
65 Jahre u. älter	72	24	20	27	143
Summe	235	184	145	265	829

4. Einführung in das Verfahren der Korrespondenzanalyse

Wir gehen von einem beliebigen Punkt aus, der sich irgendwo in einem real existierenden dreidimensionalen Raum befindet. Dieser Punkt kann anhand seiner Koordinaten (Länge, Breite und Höhe jeweils in Zentimetern) genau lokalisiert und seine Entfernung zu anderen ebenfalls im Raum befindlichen Punkten bestimmt werden. Wir können diesen Punkt, ebenso wie alle anderen Punkte, die sich in demselben Raum befinden mögen, in vektorieller Schreibweise (die Komponenten sind die Koordinaten des dazugehörigen Punktes) darstellen.

Werden alle diese Punkte in eine Matrix geschrieben, so hat diese drei Spalten, die Anzahl der Zeilen entspricht der Anzahl der Punkte. Dies ist vom formalen Aufbau her ähnlich der Tabelle 2. Der Unterschied besteht



lediglich darin, daß es sich hierbei nicht um Koordinaten von Punkten, sondern um "Altersvektoren" handelt, deren Komponenten entsprechend der jeweiligen Typenhäufigkeit determiniert sind. Dies entspricht dem Eingangsmaterial der Korrespondenzanalyse: Einer zweidimensionalen Kontingenztabelle als Matrix mit Spaltenvektoren (im Beispiel die vier verschiedenen Typen - also "Typenvektoren") und Zeilenvektoren (hier die sechs verschiedenen Altersgruppen - also "Altersvektoren").

Um zu einer grafischen Darstellung zu gelangen, sollen die Punkte, oder vielmehr die Altersausprägungen, derartig in einen zweidimensionalen Unter-
raum projiziert werden, daß die quadrierten Abstände (Distanzen) und damit auch die Informationsverluste minimal sind. Bei dieser Projektion müssen die relativen Anteile der einzelnen Variablenausprägungen (sowohl die der unabhängigen als auch die der abhängigen) berücksichtigt werden. So beträgt z.B. in der Studie der Anteil der bis 25-jährigen 11,6%, während der Anteil der 25- bis 34-jährigen 20,5% beträgt. Diese sollen entsprechend ihrer wahren Anteile (benannt als "Massen") in die Auswertung eingehen und nicht in einem Verhältnis von eins zu eins. Dies wäre aber der Fall, wenn die Randsummen der Kontingenztabelle (sie gehören nicht zu unserer Matrix der Ausgangsdaten) unberücksichtigt blieben.

Demnach besteht das Problem, daß die Punkte, die in einen niederdimensionalen Raum projiziert werden sollen, unterschiedliche Massen haben, oder anders ausgedrückt, unterschiedlich schwer sind. Zur Lösung führen wir den Schwerpunkt als Hilfsgröße ein. Wir definieren ihn als den Punkt im Raum, an welchem alle Punkte, die in die Analyse eingehen sollen, im Gleichgewicht sind. Eine anschauliche Analogie dazu bietet sich im eindimensionalen Fall bei einer Apothekerwaage, bei der die eine Seite durch Gewichte auf der anderen Seite aufgewogen wird. Was bei dieser Waage der Auflagepunkt ist, ist hier der Schwerpunkt, also der Punkt, an dem die Summe der linksdrehenden Momente gleich der Summe der rechtsdrehenden Momente ist.

Kommen wir zurück zu Tabelle 2 und relativieren vorerst alle Zeilen auf die gleiche Länge, die 1 sein soll. Die Berechnung der neuen Werte geschieht durch einfache Prozentuierung. Da die unterschiedlichen Massen von Zeilen (unterschiedliche Anteile der Altersgruppen) und der Spalten (unterschiedliche Anteile der verschiedenen Typen) berücksichtigt werden



müssen, werden auch die Randsummen jeweils ins Verhältnis zur Gesamtsumme gesetzt. Das Ergebnis dieses ersten Schrittes zur Problemlösung ist in Tabelle 3 dargestellt.

Tabelle 3: Relative Verteilung der Typen auf die Altersklassen

	macht nichts	nur eig. Stadtteil	nur City	Stadtteil und City	Summe	Masse d. Profils
bis 25 Jahre	.156	.240	.177	.427	1.0	.116
25 bis 34 Jahre	.206	.218	.188	.388	1.0	.205
35 bis 44 Jahre	.175	.277	.199	.349	1.0	.200
45 bis 54 Jahre	.266	.245	.144	.345	1.0	.168
55 bis 64 Jahre	.409	.174	.200	.217	1.0	.139
65 Jahre u. älter	.503	.168	.140	.189	1.0	.172
D_z	.283	.222	.175	.320	1.0	--

Die nun relativen Anteile der Zeilen werden als "Masse des Profils" bezeichnet. Die relativen Anteile der Spalten, welche ebenso die durchschnittlichen Anteile der Zeilen sind, "Durchschnittszeilenprofil" (D_z).

Letzteren Vektor hätten wir auch direkt aus den transponierten Vektoren der einzelnen Typen multipliziert mit dem Vektor der "Masse des Profils" (die Summe der Komponenten bliebe unberücksichtigt, da in allen Fällen gleich 1) bestimmen können, er ist der Schwerpunkt. In diesem Punkt sind somit die Informationen der relativen Spaltenanteile enthalten.

Entsprechend Tabelle 3 summieren sich die Komponenten von jedem Zeilenvektor zu Eins. Dies hat zur Folge, daß die sechs Punkte innerhalb des vierdimensionalen Raumes einen höchstens dreidimensionalen Raum aufspannen, der orthogonal zum Vektor (1,1,1,1) ist.

Die nächste Aufgabe besteht jetzt darin, diesen dreidimensionalen Raum unter Berücksichtigung der "Massen" (sowohl der Zeilen als auch der Spalten) in einen zweidimensionalen Unterraum zu projizieren. Gesucht ist eine Transformation, die die einzelnen Vektoren in einen neuen - vorerst wiederum



dreidimensionalen - Raum überführt, in dem die euklidische Metrik angenommen wird. In einem ersten Schritt sollen lediglich die relativen Spaltenanteile berücksichtigt werden. Als Lösung dieses Problems läßt sich folgende Transformation angeben (vgl. GREENACRE, 1984):

$$(1) \mathbf{a}_i^* = \mathbf{D}_z^{-1/2} \mathbf{a}_i$$

Hierbei sind die \mathbf{a}_i die Altersvektoren im alten (ursprünglichen) Raum, die \mathbf{a}_i^* die Altersvektoren im neuen Raum, und \mathbf{D}_z ist eine Matrix, in welcher in der Hauptdiagonalen die Elemente des Schwerpunktes stehen, während der Rest aus Nullen besteht. Also läßt sich die benötigte Matrix $\mathbf{d}_z^{-1/2}$ bestimmen.

Führen wir diese Transformation für die sechs Altersvektoren durch, so erhalten wir sechs neue Vektoren in einem neuen Raum unter Beibehaltung der "Masse des Profils" - also der alten "Massen" der Zeilen. Mit genau der gleichen Transformation wird auch der Schwerpunktvektor in den neuen Raum überführt, der wieder Schwerpunkt der transformierten Punktwolke ist. Als quadrierte Distanz zwischen zwei Vektoren (hier Altersvektoren) ergibt sich:

$$(2) d^2(\mathbf{a}_i, \mathbf{a}_j) = (\mathbf{a}_i - \mathbf{a}_j)^T \mathbf{D}_z^{-1} (\mathbf{a}_i - \mathbf{a}_j)$$

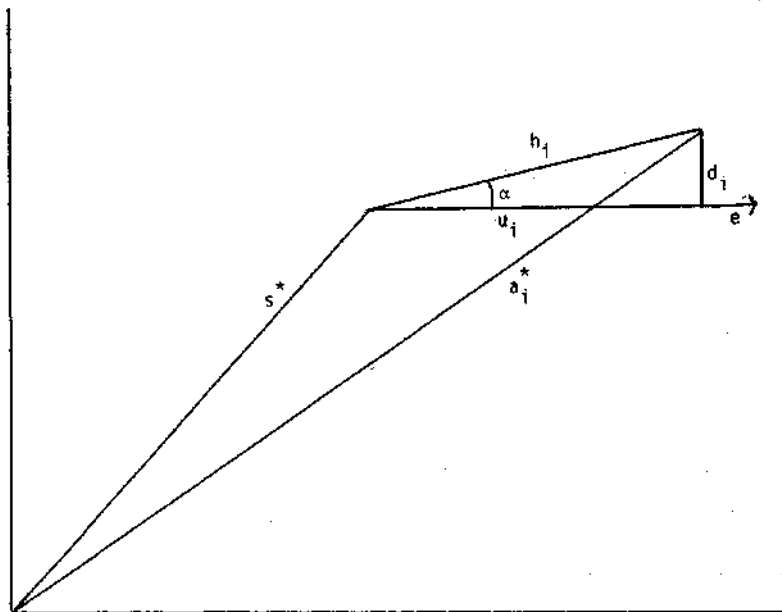
Durch die Diagonalmatrix \mathbf{D}_z^{-1} werden Unterschiede, die sich allein durch unterschiedliche Spaltensummen (hier Typenhäufigkeit) ergeben können, relativiert (vgl. GREENACRE, 1981, S. 125). Drücken wir die Distanzen mit den transformierten Vektoren aus, so erhalten wir:

$$(3) d^{2*}(\mathbf{a}_i^*, \mathbf{a}_j^*) = (\mathbf{a}_i^* - \mathbf{a}_j^*)^T (\mathbf{a}_i^* - \mathbf{a}_j^*)$$

Dies entspricht der bekannten euklidischen Metrik. Bevor wir die neuen Vektoren in einen niederdimensionalen Raum projizieren dürfen, müssen noch die relativen Anteile der Zeilen berücksichtigt werden. Die Lösung wird nach Veranschaulichung der Projektion relativ einfach sein.

Um den zur grafischen Darstellung gewünschten zweidimensionalen Unter-
raum aufspannen zu können, müssen wir zuvor, vom Schwerpunkt aus-
gehend, einen Satz paarweise orthogonal aufeinanderstehender Einheitsvek-
toren der Länge 1 bestimmen.

Abbildung 1: Grafische Darstellung der Projektion



Das Problem besteht darin, die d_i^2 zu minimieren, d.h. die quadrierten Abstände zu dem neuen Unterraum (der durch zwei Einheitsvektoren aufgespannt wird) sollen möglichst klein sein. Bekannt ist h_i als Differenz von $(a_i^* - s^*)$; aus den bekannten Informationen läßt sich auch der Einheitsvektor e (s. Abbildung 1) berechnen, worauf an späterer Stelle noch ausführlicher eingegangen wird. Der Vektor u_i entspricht der Entfernung auf dem Einheitsvektor zum Schwerpunkt, der sich durch die Projektion ergibt.



Aus der Definition des Skalarproduktes ergibt sich für

$$(4) u_i = (a_i^* - s^*)^T e, \text{ quadriert}$$

$$(5) u_i^2 = e^T (a_i^* - s^*) (a_i^* - s^*)^T e$$

Des weiteren ergibt sich aus dem Satz des Pythagoras, daß

$$(6) h_i^2 = u_i^2 + d_i^2, \text{ oder durch Summation:}$$

$$(7) \sum h_i^2 = \sum u_i^2 + \sum d_i^2$$

Dies ist gleichbedeutend mit

$$(8) \sum d_i^2 = \sum h_i^2 - \sum u_i^2$$

Die Summe der d_i^2 soll minimiert werden; dies ist wiederum erfüllt, wenn wir die Summe der u_i^2 maximieren.

Kehren wir zurück zur Berechnung der Einheitsvektoren, mit welchen der (zweidimensionale) Unterraum aufgespannt werden soll. Hierfür betrachten wir den bereits eingeführten dreidimensionalen Raum. Dieser Raum wird bestimmt durch den Schwerpunkt und die projizierten Punkte (Vektoren), die sozusagen die äußere Abgrenzung des Raumes markieren. Dieses kann durch folgenden Ausdruck beschrieben werden:

$$(9) \sum (a_i^* - s^*) (a_i^* - s^*)^T$$

An dieser Stelle kann die "Masse des Profils" - also die relativen Anteile der Zeilen (m_i) - mit in den Ausdruck einbezogen werden. Somit wurden sowohl die relativen Anteile der Spalten als auch der Zeilen berücksichtigt.

$$(10) \sum m_i (a_i^* - s^*) (a_i^* - s^*)^T$$

In dem durch Ausdruck (10) determinierten Raum soll der optimale Unterraum aufgespannt werden. Da dieser ein Maximum an Varianz erklären soll, werden die Eigenwerte benötigt, um mit diesen die Eigenvektoren bestimmen zu können. Kehren wir zur Gleichung (8) zurück. Dort haben wir ge-



sagt, daß die Summe der u_i^2 maximal sein soll. Um dies zu erreichen, müssen wir Gleichung (5) über alle i unter Berücksichtigung der relativen Anteile der Zeilen $\{m_i\}$ maximieren. Betrachten wir daher folgenden Ausdruck:

$$(11) \quad e^T \left(\sum m_i (a_i^* - s^*) (a_i^* - s^*)^T \right) e, \text{ mit } e^T e = 1$$

Dieser Ausdruck wird maximal durch geeignete Wahl von e . Dies bedeutet, daß e der Eigenvektor zum größten Eigenwert von Ausdruck (10) ist. Für die Eigenwerte erhalten wir:

$$\lambda_1 = 0,0830 \qquad \lambda_2 = 0,0035 \qquad \lambda_3 = 0,0025$$

Aufgrund der Dreidimensionalität des neuen Raumes ist λ_4 Null. Für den aus dem größten Eigenwert resultierenden Eigenvektor ergibt sich:

$$e_1 = (0,825; -0,251; -0,095; -0,497)$$

Bestimmen lassen sich nun die $u_i = (a_i^* - s^*)^T e_1$, womit wir die Projektion der "Altersvektoren" auf die erste Dimension beendet haben.

Analog gehen wir vor, um die zweite Dimension des Unterraumes zu bestimmen, indem wir den Eigenvektor für den zweiten Eigenwert ($\lambda_2 = 0,0035$) ausrechnen und mittels dessen die Projektion in die zweite (zusätzliche) Dimension vornehmen. Die "Altersvektoren" sind jetzt in einem optimalen zweidimensionalen Unterraum projiziert und lassen sich anhand der u -Werte grafisch darstellen.

Relativ einfach ist es, die erklärte Varianz der einzelnen Achsen zu berechnen. Dies geschieht, indem die einzelnen Eigenwerte durch die Summe aller Eigenwerte dividiert werden. Für die erste Achse ($\lambda_1 = 0,0830$) erhalten wir 93,2%, für die zweite noch zusätzliche 4,0% Varianzerklärung. Dies wirft für die inhaltliche Interpretation die Frage auf, ob es in diesem Fall nicht sinnvoller wäre, auf die zweite Dimension zu verzichten. Doch lassen wir sie aus didaktischen Gründen weiter bestehen.

Im zweiten Teil der Aufgabe wären die "Typenvektoren" zu betrachten. Es läßt sich zeigen, daß sich bei analogem Vorgehen (d.h. es werden rela-



tive Spaltenanteile betrachtet und der Schwerpunkt aus dem "Durchschnitts-spaltenprofil" bestimmt) genau die gleichen Eigenwerte ergeben wie die eben berechneten, da das Ausgangsmaterial in beiden Fällen identisch ist. Geändert hat sich nur der Ausgangspunkt der Betrachtung. Ferner lassen sich sowohl der Schwerpunkt als auch die neuen Einheitsvektoren aus den zuvor berechneten Werten ableiten. Zeigen ließe sich dieser Sachverhalt mit dem Theorem von ECKART und YOUNG (1936).

Um dies anschaulich zu belegen, kehren wir zu dem Beispiel mit den Punkten im Raum zurück. Jene Punkte ließen sich vektoriell mit Angaben über die Entfernung in der Länge, der Breite und der Höhe beschreiben. Diese Punkte ließen sich zurückrechnen, wenn der "Längenvektor", der "Breitenvektor" und der "Höhenvektor" bekannt wären und wir davon ausgehen können, daß die Reihenfolge der Werte unverändert geblieben ist. Es ist zulässig, beide Teilaufgaben (Projektion der Zeilen und Projektion der Spalten) in einer grafischen Darstellung zu repräsentieren und zu interpretieren (GREENACRE, 1981, p. 130).

5. Interpretation der Ergebnisse

Im folgenden Abschnitt soll etwas über die Interpretationsmöglichkeiten der Korrespondenzanalyse gesagt werden. Die Schwierigkeit liegt offensichtlich darin, daß die Abstände zwischen zwei Punkten, infolge der unterschiedlichen Varianzerklärung der Achsen und der Verzerrung durch den Dimensionsverlust, nicht mit dem Lineal gemessen und verglichen werden dürfen, - ein Problem, das nicht spezifisch für die Korrespondenzanalyse ist. Zulässig ist hingegen, die Entfernungen im relativen Vergleich zu betrachten und die Korrelationen der einzelnen Variablenausprägungen mit den (beiden) Achsen zu interpretieren. Dies kann danach geschehen,

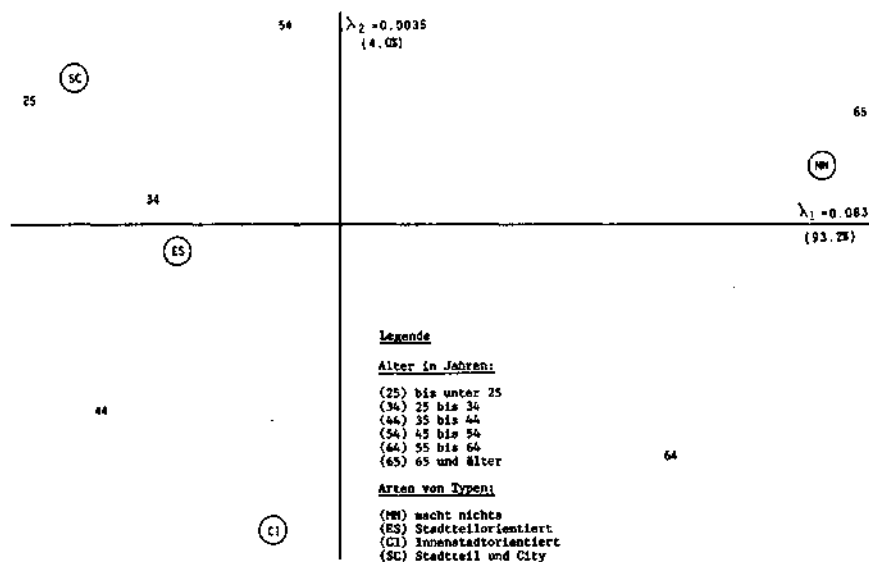
- a) in welcher Dimension die einzelnen Variablenausprägungen liegen (hier eine sehr deutliche Ähnlichkeit mit den bekannteren Verfahren der Faktorenanalyse)
- b) auf welcher Seite sie sich wie weit vom Schwerpunkt (in der grafischen Darstellung durch das Achsenkreuz gekennzeichnet) entfernt, befinden.



5.1. Bivariate Auswertung

Betrachten wir Abbildung 2, in der die Ergebnisse unseres Rechenbeispiels grafisch dargestellt sind. Wir erkennen, daß die erste Dimension, dargestellt durch die Abzisse, auf der linken Seite durch jene Gruppe von Personen gekennzeichnet ist, welche relativ viele Aktivitäten sowohl in der Innenstadt als auch im eigenen Stadtteil oder ihre Aktivitäten überwiegend im eigenen Stadtteil ausüben. Dementgegen ist auf der rechten Seite der ersten Achse jene Gruppe von Personen zu finden, die insgesamt relativ wenige Aktivitäten ausüben. Dies korrespondiert mit den Ergebnissen für die Altersgruppen. Während auf der Seite der Aktiveren die jüngeren (bis 44 Jahre) zu finden sind, so sind es auf der Seite der wenig Aktiven die älteren (ab 55 Jahre).

Abbildung 2: Grafische Darstellung der Beispielaufgabe



Die zweite Achse wird im negativen Bereich durch jene Gruppe von Personen determiniert, welche ihre Aktivitäten überwiegend in der Innenstadt ausüben. Im positiven Abschnitt dieser Achse findet sich die Gruppe von

Personen, die sich im Alter von 45 bis 54 Jahren befindet. Neben der grafischen Darstellung gibt es numerische Ergebnisse, so die Korrelationen der Variablenausprägungen mit den einzelnen Achsen - diese entsprechen dem Cosinus des Winkels (α) in Abbildung 1.

Betrachten wir diese Korrelationen der Variablenausprägungen mit den beiden Achsen, so sind diese mit der zweiten Achse mit $-,83$ (Innenstadtorientiert) und $,71$ (45 bis 54 Jahre) im Vergleich zu den Korrelationen mit der ersten Achse von $-,999$ (Stadtteilorientiert), $-,97$ (25 bis 34 Jahre) und $,996$ (über 64 Jahre) relativ niedrig. Zudem sollte die zweite Achse infolge ihrer geringen Erklärungskraft (4,0%) nicht überinterpretiert werden.

Diese scheinbar trivialen Ergebnisse decken sich zwar mit den Ergebnissen bisheriger Forschung (oder auch dem Alltagswissen), doch rechtfertigen sie an dieser Stelle noch keineswegs den Aufwand einer besonderen Analyse, zudem es gerade für die bivariate Auswertung eine Vielzahl von geeigneteren Methoden gibt. Wie zu Anfang aber gesagt, wollen wir uns gerade nicht auf den bivariaten Fall beschränken, sondern die Korrespondenzanalyse zur multivariaten Auswertung verwenden.

5.2. Multivariate Auswertung

In Analogie zu GREENACRE (1984, S. 76) haben wir unser Beispiel um 12 zusätzliche Variablen mit nun insgesamt 40 Variablenausprägungen (siehe Tabelle 4) erweitert. Als Eingabeinformation benötigen wir lediglich mehrere Kontingenztabellen, in denen die einzelnen Variablen mit der zu erklärenden Variable "Aktivitätstyp" kreuztabelliert wurden. Diese einzelnen Tabellen wurden untereinander geschrieben (vgl. BLASIUS und ROHLINGER, 1987, BLASIUS, 1987, S. 185) und bilden in dieser Darstellungsform das Ausgangsmaterial der multivariaten Analyse. Die rechnerische Vorgehensweise war identisch mit der eben vorgestellten, die grafische Darstellung der Ergebnisse gibt Abbildung 3 wieder.

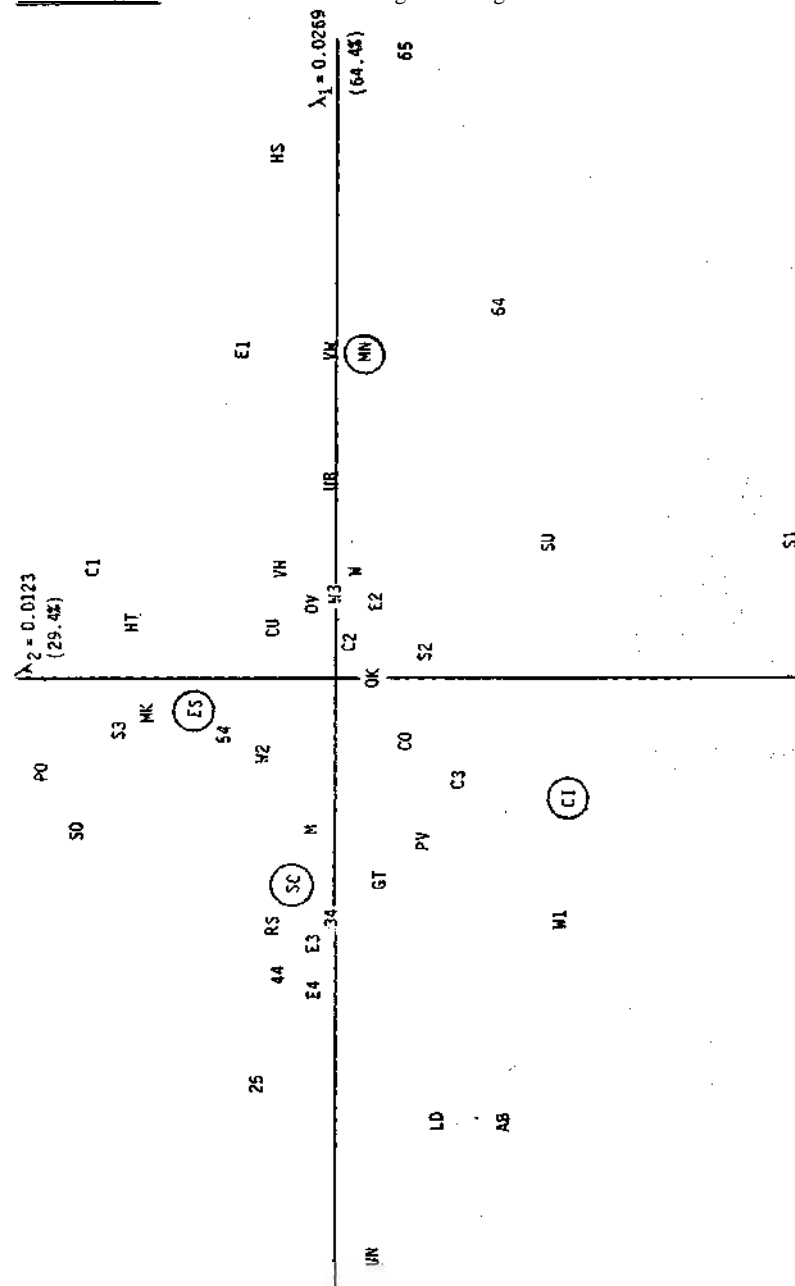
Die horizontale Achse erklärt 64,4% der Varianz; sie ließe sich als "Dimension der sozio-demografischen Merkmale" der Personen beschreiben. Während sich auf der rechten Seite vom Schwerpunkt jene Personengruppen befinden, die schon älter sind, über einen Hauptschulabschluß verfügen, verwitwet sind und/oder nur ein geringes Einkommen haben, befinden sich auf der



Tabelle 4: Darstellung der numerischen Ergebnisse

Verwendete Variablen	Gesamtmodell			1. Achse			2. Achse		
	SqKor	Nasse	Trg.	Lage	qKor	Trg.	Lage	qKor	Trg.
Familienstand:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
verheiratet (VH)	1000	38	10	90	766	11	50	234	8
verwitwet (VW)	957	18	29	256	956	43	8	1	0
ledig (LD)	984	22	70	-343	911	99	-97	73	17
Kinder unter 10 Jahre:									
nein (OK)	852	65	2	7	37	0	-33	815	6
ja (MK)	815	13	12	-34	32	1	175	783	32
Bildungsabschluß:									
Hauptschule (HS)	969	32	125	395	957	186	45	12	5
Realschule (RS)	1000	23	22	-193	930	32	53	70	5
Abitur (AB)	985	15	50	-332	793	62	-163	192	33
Universität (UN)	901	8	41	-443	897	57	-30	4	1
Geschlecht:									
männlich (M)	806	36	12	-105	788	15	16	18	1
weiblich (W)	810	42	11	92	791	13	-13	19	1
Alter in Jahren:									
bis unter 25 (25)	987	9	20	-297	939	30	67	48	3
25 bis 34 (34)	967	16	14	-187	967	21	2	0	0
35 bis 44 (44)	789	16	23	-214	756	27	45	33	3
45 bis 54 (54)	981	13	4	-35	113	1	99	868	10
55 bis 64 (64)	999	11	26	282	776	32	-150	223	20
65 und älter (65)	985	13	79	488	969	119	-62	16	4
Beschäftigungsverhältnis:									
ganztags (GT)	991	36	25	-162	917	35	-45	74	6
halbtags (HT)	847	6	7	39	34	0	194	813	19
unbeschäftigt (UB)	966	36	23	160	965	35	6	1	0
Haushaltseinkommen:									
bis 1499,-DM (E1)	1000	13	23	252	892	31	88	108	8
1500,- bis 2499,-DM (E2)	930	23	3	61	685	3	-35	245	2
2500,- bis 3999,-DM (E3)	978	20	21	-204	964	31	25	14	1
4000,-DM und mehr (E4)	891	7	11	-239	889	15	12	2	0
Verkehrsmittel i.d.City:									
öffentlich (OV)	577	50	8	62	552	7	13	25	1
privat (PV)	661	23	20	-127	456	14	-85	205	14
privat u. öffentlich (PO)	997	3	6	-68	57	1	279	940	20
Einstellung zum Stadtteil:									
negativ (S1)	962	13	61	97	48	4	-423	914	188
mittel (S2)	782	25	6	23	54	0	-84	728	14
positiv (S3)	1000	41	35	-41	50	3	185	950	113
Einstellung zur Innenstadt:									
negativ (C1)	993	18	24	91	149	6	215	844	69
mittel (C2)	171	28	2	24	158	1	-6	13	0
positiv (C3)	938	32	15	-68	243	6	-116	695	36
Bew. Ausstattung Stadtteil:									
unterhalb des Medians (SU)	997	43	48	100	214	16	-190	783	128
oberhalb des Medians (SO)	997	35	59	-120	208	19	236	789	159
Bew. Ausstattung Innenstadt:									
unterhalb des Medians (CU)	425	38	13	45	149	3	61	276	12
oberhalb des Medians (CO)	422	40	12	-40	137	3	-58	285	11
Wohndauer im Stadtteil:									
bis 2 Jahre (W1)	998	10	21	-186	417	13	-220	581	41
3 bis 5 Jahre (W2)	649	11	3	-61	303	2	67	346	4
über 5 Jahre (W3)	982	57	4	51	783	5	26	199	5
	1000	1000			1000				1000
Arten von Typen:									
nicht orientiert (NN)	993	283	431	251	990	661	-12	3	4
Stadtteilorientiert (ES)	784	221	127	-26	30	6	135	754	326
Innenstadtorientiert (CI)	966	176	232	-91	153	55	-211	813	639
Stadtteil und City (SC)	892	320	211	-152	849	278	35	43	31
	1000	1000			1000				1000

Abbildung 3: Grafische Darstellung der Ergebnisse



Legende zu Abbildung 3Familienstand:

- (VH) verheiratet
(VW) verwitwet
(LD) ledig

Kinder unter 10 Jahre:

- (OK) nein
(MK) ja

Bildungsabschluß:

- (HS) Hauptschule
(RS) Realschule
(AB) Abitur
(UN) Universität

Geschlecht:

- (M) männlich
(W) weiblich

Alter in Jahren:

- (25) bis unter 25
(34) 25 bis 34
(44) 35 bis 44
(54) 45 bis 54
(64) 55 bis 64
(65) 65 und älter

Beschäftigungsverhältnis

- (GT) ganztags
(HT) halbtags
(UB) unbeschäftigt

Haushaltseinkommen:

- (E1) bis 1499,- DM
(E2) 1500,- bis 2499,- DM
(E3) 2500,- bis 3999,- DM
(E4) 4000,- DM und mehr

Verkehrsmittel i.d. City:

- (OV) öffentlich
(PV) privat
(PO) privat u. öffentlich

Einstellung zum Stadtteil:

- (S1) negativ
(S2) mittel
(S3) positiv

Einstellung zur Innenstadt:

- (C1) negativ
(C2) mittel
(C3) positiv

Bew. Ausstattung, Stadtteil:

- (SU) unterhalb des Medians
(SO) oberhalb des Medians

Bew. Ausstattung, Innenstadt:

- (CU) unterhalb des Medians
(CO) oberhalb des Medians

Wohndauer im Stadtteil:

- (W1) bis 2 Jahre
(W2) 3 bis 5 Jahre
(W3) über 5 Jahre

Arten von Typen:

- (MN) macht nichts
(ES) Stadtteilorientiert
(CI) Innenstadtorientiert
(SC) Stadtteil und City



linken Seite diejenigen, die über eine höhere Schulbildung und ein höheres Einkommen verfügen und/oder einer jüngeren Altersklasse angehören. Auch haben wir hier die Ganztagsbeschäftigten, während die Gruppe der Unbeschäftigten in der Nähe der älteren und Beziehher geringerer Einkommen liegt. Dies korrespondiert mit den unterschiedlichen Typen. Während diejenigen, die sowohl in der Innenstadt als auch im eigenen Stadtteil relativ viele Aktivitäten ausüben, sich auf der Seite der besser Verdienenden befinden, sind jene, die relativ wenig Aktivitäten ausüben, auf der Seite der Älteren, Verwitweten und der Unbeschäftigten.

Die zweite Achse erklärt 29,4% der Varianz, sie kann als "Dimension der Einstellungen" interpretiert werden. Während sich hier auf der einen Seite diejenigen befinden, welche zwar eine positive Einstellung zum eigenen Stadtteil, aber eine negative zur Innenstadt haben, sind auf der anderen Seite die Verhältnisse umgekehrt. Dies stimmt mit den ausgeübten Aktivitäten überein, so daß wir sagen können, daß die Einstellung zum eigenen Stadtteil und zur Innenstadt sehr eng mit dem Ort der ausgeübten Aktivitäten zusammenhängt. Interessant ist auch, daß bei Befragten mit kleinen Kindern scheinbar eine deutliche Stadtteilorientierung gegeben ist, denn dieser Personenkreis befindet sich in unmittelbarer Nähe zu dem als "Stadtteiltyp" definierten Personenkreis.

Dem entgegen läßt sich jener Personenkreis, der den Schwerpunkt seiner Aktivitäten in der Innenstadt hat, nicht mit Hilfe sozio-demografischer Merkmale der Befragten beschreiben. Aufgrund der hier einbezogenen Variablen ist zu schließen, daß es so etwas wie den typischen Innenstadtbesucher nicht gibt, vielmehr die Innenstadt über alle Personenkreise hinweg genutzt oder nicht genutzt wird.

Um zu einer weitergehenden Interpretation der Daten zu gelangen, bedarf es der numerischen Ergebnisse: sie sind in Tabelle 4 dargestellt. Anhand dieser Daten ist deutlich zu erkennen, daß die erste Achse bei den unabhängigen Variablen hauptsächlich mit den sozio-demografischen Merkmalen der Person korreliert, während keinerlei Einstellungsmerkmale darauf laden.

In der ersten Spalte der Tabelle 4 (SqKor) ist die Varianz dargestellt, die durch die ersten beiden Achsen erklärt wird. Hier wird z.B. ersichtlich, daß die Gruppe der Verheirateten und die Gruppe der Realschüler durch



diese beiden Achsen vollständig, d.h. zu 100%, erklärt werden, während bei der Gruppe derjenigen, die über einen Universitätsabschluß verfügen, noch nahezu 10% der Varianz durch die dritte (und letzte) Achse determiniert werden. In den Spalten 5 und 8 (qKor) sind die quadrierten Korrelationen mit den beiden dargestellten Achsen notiert. Ihre Summe entspricht der ersten Spalte.

In der zweiten Spalte (Masse) befindet sich die relative Masse der einzelnen Variablenausprägungen bzw. der vier verschiedenen Typen, diese ist für Zeilen und Spalten getrennt auf 1000 standardisiert. Hier ist z.B. zu erkennen, daß die Gruppe der Befragten mit Kindern unter 10 Jahren im Haushalt in einem Verhältnis von 13 zu 65 zu der Gruppe von Befragten steht, in denen keine Kinder unter 10 Jahren im Haushalt leben. In dieser Spalte finden auch fehlende Werte Berücksichtigung. So hat die Gesamtmasse des Einkommens nur einen Anteil von 6,3%, während z.B. das Geschlecht einen Anteil von 7,8% hat. (Es sind die Anteile der einzelnen Variablenausprägungen jeweils zu addieren.)

Das "Trägheitsgewicht" (Trg.) gibt an, wie stark das Modell (Spalte 3) bzw. die einzelnen Achsen (Spalte 6 und 9) von den einzelnen Variablenausprägungen bzw. den "Typen" determiniert wird. Es ist ähnlich unserer Apothekerwaage, wo ein größeres Gewicht (hier als Anzahl von Personen) durch eine entsprechend größere Entfernung eines Gewichtes auf der anderen Seite vom Auflagepunkt ausgeglichen werden kann. Berechnet wird das Trägheitsgewicht aus der "Masse", multipliziert mit dem Quadrat der Entfernung vom Schwerpunkt auf der jeweiligen Achse. Diese Trägheitsgewichte addieren sich über die Variablenausprägungen - und davon getrennt auch über die vier "Typen" zu 1000 (entspricht 100%). Während in den Spalten qKor und SqKor die erklärte Varianz der Variablenausprägungen durch die Achsen angegeben wird, befindet sich in den Spalten 6 und 9 die erklärte Varianz der Achsen durch die einzelnen Variablenausprägungen. In der dritten Spalte wurden diese erklärten Varianzen mit der Erklärungskraft der einzelnen Achsen gewichtet, somit ist hier angegeben, wie stark das gesamte Modell von den einzelnen Variablenausprägungen - für Zeilen und Spalten getrennt - determiniert wurde.

Die Spalten 4 und 7 geben die Lage der Variablenausprägungen und der "Typen" auf den einzelnen Achsen an. Hier ist erkennbar, auf welcher



Seite der Achse sich welche Variablenausprägungen wie weit vom Schwerpunkt entfernt befinden.

Deutlich werden unsere vier "Typen" durch die ersten beiden Achsen diskriminiert, eine Zuordnung zu den jeweiligen Merkmalen erfolgt anhand von Tabelle 4. So befindet sich jene Gruppe von Personen, die sowohl in der Innenstadt als auch im eigenen Stadtteil relativ wenige Aktivitäten ausübt, auf der rechten Seite der ersten Achse mit einem durch diese erklärten Varianzanteil von 99%. Wählen wir als "Schwellenwert" (die Festsetzung des Wertes erfolgt analog zur Zuordnung von Variablen zu Faktoren bei der Hauptkomponentenanalyse) für die Zuordnung einer Variablenausprägung zu einem Achsenabschnitt 700 (d.h. 70% der Varianz sollen durch diese Achse erklärt sein), so lassen sich folgende Variablenausprägungen (Gruppen von Personen) dem positiven Bereich der ersten Achse (positives Vorzeichen in der Spalte "Lage") - und damit der Gruppe von Personen, die relativ wenige Aktivitäten in der Innenstadt und im eigenen Stadtteil ausübt - zuordnen: die Verheirateten, die Verwitweten, diejenigen mit "Hauptschulabschluß", die Frauen, die über 54-jährigen, die Unbeschäftigten, die Bezieher geringer Einkommen und diejenigen, die schon über 5 Jahre in ihrem derzeitigen Stadtteil wohnen.

Trotz ähnlich hoher Anteile erklärter Varianz durch die erste Achse haben nicht alle Variablenausprägungen für die Modelldetermination die gleiche Bedeutung. So "laden" zwar die Gruppe der Verwitweten und die Gruppe der Verheirateten im positiven Bereich der ersten Achse - und sind damit beide positiv mit der Gruppe der wenig Aktiven korreliert -, doch hat die Gruppe der Verwitweten den fast vierfachen Anteil an der Determination wie die Gruppe der Verheirateten (vgl. Spalte 6, Tabelle 4), obwohl sie nicht einmal halb so viele Personen (vgl. Spalte 2) umfaßt. Inhaltlich bedeutet dies, daß das Ausführen von nur wenigen Aktivitäten für Verwitwete wesentlich charakteristischer ist als für Verheiratete.

Im positiven Bereich der ersten Achse - also negativ mit den eben genannten Variablenausprägungen korreliert - befindet sich die Gruppe von Personen, die sowohl in der Innenstadt als auch im eigenen Stadtteil überdurchschnittlich viele Aktivitäten ausübt. Diese Personengruppe läßt sich beschreiben als überdurchschnittlich häufig ledig, als diejenigen, die über einen höheren Schulabschluß verfügen (mindestens Realschule), der Männer,



der unter 45-jährigen, der Ganztagsbeschäftigten, sowie als Bezieher höherer Einkommen (2500,- und mehr).

Während die erste Achse durch den Gegensatz der Gruppen der überdurchschnittlich zu den unterdurchschnittlich Aktiven sowohl in der Innenstadt als auch im eigenen Stadtteil gekennzeichnet ist, ist die zweite Achse - die entsprechend dem Modell der Korrespondenzanalyse unkorreliert mit der ersten und den anderen Achsen ist - durch die Gruppe der Stadtteil- vs. der Gruppe der Innenstadtorientierten determiniert. Die Gruppe von Personen, die überdurchschnittlich viele Aktivitäten im eigenen Stadtteil hat, läßt sich durch folgende Merkmale kennzeichnen: das Vorhandensein von Kindern unter 10 Jahren im Haushalt, der 45- bis 54-jährigen, Halbtagsbeschäftigung, die Benutzung von privaten und öffentlichen Verkehrsmitteln, eine positive Einstellung zum eigenen Stadtteil und eine negative zur Innenstadt, sowie eine überdurchschnittliche Bewertung des eigenen Stadtteils. (Vorgegeben wurden je 12 Einrichtungen für die Innenstadt und den eigenen Stadtteil, für die die Befragten eine Note zwischen "sehr gut" und "schlecht" geben sollten.) Im negativen Bereich der zweiten Achse befinden sich die Innenstadtorientierten. Sie lassen sich charakterisieren als Gruppen von Personen, bei denen keine Kinder unter 10 Jahren im Haushalt leben, die eine negative oder eine mittlere Einstellung zum eigenen Stadtteil haben und die dessen Ausstattung auch unterdurchschnittlich bewerten. Ferner läßt sich hier die Gruppe von Personen zuordnen, die der Innenstadt gegenüber positiv eingestellt sind.

Auf der dritten, in Tabelle 4 nicht dargestellten Achse (erklärte Varianz: 6,2%), "läßt" die Merkmalsausprägung "mittlere Einstellung zur Innenstadt" und - etwas schwächer (unterhalb des von uns gewählten "Schwellenwertes", aber noch mit über 50% erklärter Varianz durch diese Achse) - die Bewertung der Ausstattung der Innenstadt mit ihren zwei Ausprägungen.

Da keine der sozio-demografischen Variablenausprägungen mit dem Merkmal "Innenstadt-orientiert" positiv korreliert, veranlaßt uns dies zu dem Schluß, daß es so etwas wie den typischen Innenstadtbewohner nicht gibt, die Innenstadt vielmehr von nahezu allen Personenkreisen genutzt bzw. nicht genutzt wird. Die Ursache liegt vermutlich darin, daß die Innenstadt jeder Personen-



gruppe etwas für sie Spezifisches (Schaufensterbummel, Kino, Konzert, Oper) bieten kann. Beim Aggregieren über die verschiedenen Tätigkeiten gehen die einzelnen Informationen (wer welche Tätigkeit ausübt) verloren. In dieser Arbeit sollte jedoch nicht geprüft werden, ob sich die Innenstadtbesucher hinsichtlich ihrer Aktivitäten unterscheiden.

6. Fazit

Es sollte gezeigt werden, daß die Korrespondenzanalyse ein sehr gutes Instrument zur multivariaten Auswertung qualitativer Daten ist. Hypothesen wurden getestet, für die bislang ein geeignetes Verfahren fehlte. Die universelle Anwendbarkeit der Korrespondenzanalyse sollte auch Grund sein zu prüfen, ob in Anwendungsbeispielen wie dem hier vorgestellten nicht völlig auf Verfahren wie Faktoren-, Cluster- oder Diskriminanzanalyse verzichtet werden kann und muß. Die hier dargestellten Ergebnisse sollten ermutigen, weiter über das Verfahren nachzudenken.

Anmerkung

- 1) Ein von H. ROHLINGER (Zentralarchiv) und mir geschriebenes Programm ist gegen Kostenerstattung erhältlich. Es handelt sich hierbei um ein Lehrprogramm, das in der matrixorientierten Sprache SAS-PROC MATRIX geschrieben wurde. Für eine zufriedenstellende grafische Darstellung wird SAS-GRAPH in der SAS-Version 5.16 benötigt.

Literatur

BENZECRI, J.-P., 1963: Course de Linguistique Mathématique. Rennes: Université de Rennes, France.

BLASIUS, J., 1987: Korrespondenzanalyse - ein multivariates Verfahren zur Analyse qualitativer Daten. Historische Sozialforschung 42-43: 172-189.

BLASIUS, J./ROHLINGER, H., 1987: KORRES - ein Programm zur Korrespondenzanalyse. Vortragspapier zur Statistik-Software Konferenz in Heidelberg vom 23.3.-26.3. 1987. Köln: Zentralarchiv für empirische Sozialforschung.

DANGSCHAT, J./BLASIUS, J., 1987: Social and Spatial Disparities in Warsaw 1978: An Application of Correspondence Analysis to a 'Socialist City'. Urban Studies 24: 173-191.

ECKART, C./YOUNG, G., 1936: The Approximation of one Matrix by Another of Lower Rank. Psychometrika 1: 211-218.



ESCOFIER-CORDIER, B., 1965: I/Analyse des Correspondances. Unveröffentlichte Dissertation: Université de Rennes, France.

FRIEDRICHS, J./GOODMAN, A.C. et al., 1987: The Changing Downtown. A comparative Study of Baltimore and Hamburg. Berlin und New York: de Gruyter.

GREENACRE, M., 1981: Practical Correspondence Analysis. S. 119-146 in: V. BARNETT (Hrsg.), Interpreting Multivariate Data. Chichester, U.K.: Wiley.

GREENACRE, M., 1984: Theory and Applications of Correspondence Analysis. London: Academic Press.

GUTTMAN, L., 1941: The Quantification of a Class of Attributes: A Theory and Method of Scale Construction. S. 319-348 in: P. HORST (Hrsg.), The Prediction of Personal Adjustment. New York: Social Science Research Council.

GUTTMAN, L., 1950: The Principal Components of Scale Analysis. In: S.A. STOUFFER/L. GUTTMAN/E.A. SUCHMAN/P.F. LAZARSFELD/S.A. STAR/J.A. CLAUSEN (Hrsg.), Measurement and Prediction. Princeton, N.J.: Princeton University Press, S. 312-361.

HAYASHI, C., 1950: On the Quantification of Qualitative Data from the Mathematico - Statistical Point of View. Annals of the Institute of Statistical Mathematics 2: 35-47.

HAYASHI, C., 1952: On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of View. Annals of the Institute of Statistical Mathematics 3: 69-98.

HAYASHI, C., 1954: Multidimensional Quantification - with the Applications to Analysis of Social Phenomena. In: Annals of the Institute of Statistical Mathematics 5: 231-245.

HILL, O., 1973: Reciprocal Averaging: An Eigenvector Method of Ordination. Journal of Ecology 61: 237-251.

HIRSCHFELD, H.O., 1935: A Connection between Correlation and Contingency. Proceedings of the Cambridge Philosophical Society 31: 520-524.

NISHISATO, S., 1978: Optimal Scaling of Paired Comparisons and Rank Order Data. An Alternative to Guttman's Formulation. Psychometrika 43: 263-271.

NISHISATO, S., 1980: Analysis of Categorical Data: Dual Scaling and its Applications. Toronto: University of Toronto Press.

Jörg Blasius
Forschungsstelle für vergleichende Stadtforschung, Universität Hamburg
jetzt:
Seminar für Soziologie, Universität zu Köln